Adjacency relationships forced by a degree sequence

Michael D. Barrus
Department of Mathematics
University of Rhode Island
Kingston, RI 02881
barrus@uri.edu

August 4, 2015

Abstract

There are typically several nonisomorphic graphs having a given degree sequence, and for any two degree sequence terms it is often possible to find a realization in which the corresponding vertices are adjacent and one in which they are not. We provide necessary and sufficient conditions for two vertices to be adjacent (or nonadjacent) in every realization of the degree sequence. These conditions generalize degree sequence and structural characterizations of the threshold graphs, in which every adjacency relationship is forcibly determined by the degree sequence. We further show that degree sequences for which adjacency relationships are forced form an upward-closed set in the dominance order on graphic partitions of an even integer.

1 Introduction

A fundamental goal of the study of graph degree sequences is to identify properties that must be shared by all graphs having the same degree sequence. In this paper we address one of the simplest of graph properties: whether two given vertices are adjacent.

Most degree sequences d are shared by multiple distinct graphs. We call these graphs the *realizations* of d. In this paper we consider only labeled graphs, that is, we distinguish between realizations having distinct edge sets, even if these realizations are isomorphic. Throughout the paper we will consider a degree sequence $d = (d_1, \ldots, d_n)$ and all realizations of d with vertex set $V = \{1, \ldots, n\}$ (we denote such a range of natural numbers by $[n]$) that satisfy the condition that each vertex i has the corresponding degree d_i. We will assume in each case, unless otherwise stated, that $d_1 \geq \cdots \geq d_n$.

For only one type of degree sequence are all the adjacency relationships in a realization completely determined. These are the *threshold sequences*, those sequences having only one realization. *Threshold graphs*, the graphs realizing threshold sequences, were introduced (via an equivalent definition) by Chvátal and Hammer in [4, 5], as well as by many other authors independently. These graphs have a number of remarkable properties; see the monograph [11] for a survey and bibliography. We will refer to several of these properties in the course of the paper.

On the other end of the spectrum from the threshold sequences, many degree sequences have the property that *any* fixed pair of vertices may be adjacent in one realization and nonadjacent in another; such is the case, for example, with $(1, 1, 1, 1)$. For still other sequences, some adjacency relationships are determined while others are not; notice that in the two realizations of $(2, 2, 1, 1, 0)$ the vertices of degree 2 must be adjacent, the vertices of degree 1 must be nonadjacent, and the vertex of degree 0 cannot be adjacent to anything, while a fixed vertex of degree 1 may or may not be adjacent to a fixed vertex of degree 2.

Suppose that d is an arbitrary degree sequence. We classify pairs $\{i, j\}$ of vertices from V as follows. If i and j are adjacent in every realization of d, we say that $\{i, j\}$ is a *forced edge*. If i and j are adjacent in no realization of d, then $\{i, j\}$ is a *forced non-edge*. Vertices in a forced edge or forced non-edge are *forcibly*

adjacent or *forcibly nonadjacent*, respectively. If $\{i, j\}$ is either a forced edge or forced non-edge, we call it a *forced pair*; otherwise, it is *unforced*. By definition, in threshold graphs every pair of vertices is forced.

In this paper we characterize the forced pairs for general degree sequences. We present conditions that allow us to recognize these pairs from the degree sequence and describe the structure they as a set create in any realization of the degree sequence.

As an alternative viewpoint, given a degree sequence d, we may define the *intersection envelope graph* $I(d)$ (respectively, *union envelope graph* $U(d)$) to be the graph with vertex set $[n]$ whose edge set is the intersection (resp., union) of the edge sets of all realizations of d. The forced edges of d are precisely the edges of $I(d)$, and the forced non-edges of d are precisely the non-edges of $U(d)$. As we will see, $I(d)$ and $U(d)$ are threshold graphs, and our results allow us to describe these graphs.

One particular property of threshold sequences contextualized by a study of forced pairs is the location of these sequences in the dominance (majorization) order on degree sequences having the same sum. Threshold sequences comprise the maximal elements in this order, and we show that as a collection, degree sequences with forced pairs majorize degree sequences having no forced pairs.

The structure of the paper is as follows: In Section 2 we provide necessary and sufficient conditions on a degree sequence for a pair $\{i, j\}$ to be a forced edge or forced non-edge among realizations of a degree sequence d. We then give an alternative degree sequence characterization in terms of Erdős–Gallai differences, which we introduce. In Section 3 we study the overall structure of forced pairs in a graph, describing the envelope graphs $I(d)$ and $U(d)$. Finally, in Section 4 we present properties of forced pairs in the context of the dominance order on degree sequences.

Throughout the paper all graphs are assumed to be simple and finite. We use $V(H)$ to denote the vertex set of a graph H. A list of nonnegative integers is *graphic* if it is the degree sequence of some graph. A *clique* (respectively, *independent set*) is a set of pairwise adjacent (nonadjacent) vertices.

2 Degree sequence conditions for forced pairs

We begin with a straightforward test for determining whether a pair of vertices is forced.

Theorem 2.1. *Given the degree sequence $d = (d_1, \dots, d_n)$ and vertex set $[n]$, let i, j be distinct elements of $[n]$ such that $i < j$. The pair $\{i, j\}$ is a forced edge if and only if the sequence*

$$d^+(i, j) = (d_1, \dots, d_{i-1}, d_i + 1, d_{i+1}, \dots, d_{j-1}, d_j + 1, d_{j+1}, \dots, d_n)$$

is not graphic. The pair $\{i, j\}$ is a forced non-edge if and only if the sequence

$$d^-(i, j) = (d_1, \dots, d_{i-1}, d_i - 1, d_{i+1}, \dots, d_{j-1}, d_j - 1, d_{j+1}, \dots, d_n)$$

is not graphic.

Before proving this theorem, we introduce some notation. Given a degree sequence π of length n, let $\overline{\pi}$ denote the degree sequence of the complement of a realization of π, i.e., $\overline{\pi} = (n - 1 - d_n, \dots, n - 1 - d_1)$; we call $\overline{\pi}$ the complementary degree sequence of π. Note that π is also the complementary degree sequence of $\overline{\pi}$.

Proof. We begin by proving the contrapositives of the statements in the first equivalence. Suppose first that $\{i, j\}$ is not a forced edge for d. There must exist a realization G of d in which i and j are not adjacent. The graph H formed by adding edge ij to G has degree sequence $d^+(i, j)$, so $d^+(i, j)$ is graphic.

Suppose now that $d^+(i, j)$ is graphic, and let H be a realization. Suppose also that G is a realization of d. If $\{i, j\}$ is not an edge of G, then it is not a forced edge for d. Furthermore, if $\{i, j\}$ is an edge of H, then removing that edge produces a realization of d with no edge between i and j, so once again $\{i, j\}$ is not a forced edge. Suppose now that ij is an edge of G but not of H. Let J be the symmetric difference of G and H, that is, the graph on $[n]$ having as its edges all edges belonging to exactly one of G and H. Color each edge in J red if it is an edge of G and blue if it is an edge of H. Since the degree of any vertex in $[n]$

other than i and j is the same in both G and H, there is an equal number of red and blue edges meeting at such a vertex. For all such vertices, partition the incident edges into pairs that each contain a red and a blue edge. Now vertices i and j each are incident with one more blue edge than red; fix a vertex v such that iv is a blue edge in J and partition the other edges incident with i into pairs containing a red and a blue edge. Do the same thing for the edges incident with j other than a fixed blue edge jw. We now find a path from i to j in J whose edges alternate between blue and red. Note that iv is a blue edge, and that this edge is paired with a red edge incident with v, which is in turn paired with a blue edge at its other endpoint, and so on. Since each edge of J other than iv and jw is paired with a unique edge of the opposite color at each of its endpoints, the path beginning with iv must continue until it terminates with edge wj. Now let $v_0, v_1, \dots, v_\ell$ be the vertices encountered on this path, in order, so that $v_0 = i$, $v_1 = v$, $v_{\ell-1} = w$, and $v_\ell = j$. The graph G contains edges $v_1v_2, v_3v_4, \dots, v_{\ell-2}v_{\ell-1}$ and $v_\ell v_0$ and non-edges $v_0v_1, v_2v_3, \dots, v_{\ell-1}v_\ell$. Deleting these edges and adding the non-edges as new edges creates a realization of d where i and j are not adjacent, so once again $\{i, j\}$ is not a forced edge for d.

Since $\{i, j\}$ is an edge in a realization of π if and only if it is not an edge in a realization of $\overline{\pi}$, the pair $\{i, j\}$ is a forced non-edge for d if and only if it is a forced edge for $\overline{d}$, which is equivalent by the preceding paragraph to having $\overline{d}^+(i, j)$ not be graphic. Since a list π of integers is a degree sequence if and only if $\overline{\pi}$ is a degree sequence, and we can easily verify that $d^-(i, j) = \overline{\overline{d}^+(i, j)}$, the pair $\{i, j\}$ is a forced non-edge in G if and only if $d^-(i, j)$ is not a graphic sequence. □

By combining Theorem 2.1 with a test for graphicality we may find alternate characterizations of forced pairs. Here we will use the well known Erdős–Gallai criteria [6] with a simplification due to Hammer, Ibaraki, and Simeone [8, 9]). For any integer sequence $\pi = (\pi_1, \dots, \pi_n)$, define $m(\pi) = \max\{i : \pi_i \geq i - 1\}$.

Theorem 2.2 ([6, 8, 9])**.** *A list $\pi = (\pi_1, \dots, \pi_n)$ of nonnegative integers in descending order is graphic if and only if $\sum_k \pi_k$ is even and*

$$\sum_{\ell \leq k} \pi_\ell \leq k(k-1) + \sum_{\ell > k} \min\{k, d_\ell\}$$

for each $k \in \{1, \dots, m(\pi)\}$.

For each $k \in [n]$, let $\mathrm{LHS}_k(\pi) = \sum_{\ell \leq k} \pi_\ell$ and $\mathrm{RHS}_k(\pi) = k(k-1) + \sum_{\ell > k} \min\{k, d_\ell\}$. We now define the *kth Erdős–Gallai difference* $\Delta_k(\pi)$ by

$$\Delta_k(\pi) = \mathrm{RHS}_k(\pi) - \mathrm{LHS}_k(\pi).$$

Note that an integer sequence with even sum is graphic if and only if these differences are all nonnegative.

Theorem 2.3. *Let $d = (d_1, \dots, d_n)$ be a graphic list, and let i, j be integers such that $1 \leq i < j \leq n$. The pair $\{i, j\}$ is a forced edge for d if and only if there exists k such that $1 \leq k \leq n$ and one of the following is true:*

(1) $\Delta_k(d) \leq 1$ *and* $j \leq k$.

(2) $\Delta_k(d) = 0$*;* $i \leq k < j$*; and* $k \leq d_j$.

The pair $\{i, j\}$ is a forced non-edge for d if and only if there exists k such that $1 \leq k \leq n$ and one of the following is true:

(3) $\Delta_k(d) \leq 1$ *and* $d_i < k < i$.

(4) $\Delta_k(d) = 0$*;* $k < i$*; and* $d_j \leq k \leq d_i$.

Proof. By Theorems 2.1 and 2.2, $\{i, j\}$ is a forced edge if and only if there exists an integer k such that $\Delta_k\left(d^+(i, j)\right) < 0$. We prove that this happens if and only if condition (1) or condition (2) holds. Let k be an arbitrary element of $[n]$.

Case: $k < i$. In this case neither condition (1) nor condition (2) holds. Furthermore, $\mathrm{LHS}_k(d^+(i,j)) = \mathrm{LHS}_k(d) \le \mathrm{RHS}_k(d) \le \mathrm{RHS}_k(d^+(i,j))$, so $\Delta_k\left(d^+(i,j)\right) \ge 0$.

Case: $j \le k$. Here condition (2) does not hold. Since $\mathrm{RHS}_k(d^+(i,j)) = \mathrm{RHS}_k(d)$ and $\mathrm{LHS}_k(d^+(i,j)) \le \mathrm{LHS}_k(d) + 2$, we see that $\Delta_k\left(d^+(i,j)\right) < 0$ if and only if $\Delta_k(d) \le 1$, which is equivalent to condition (1).

Case: $i \le k < j$. Note that condition (1) cannot hold in this case. Since $i \le k$, we have $\mathrm{LHS}_k(d^+(i,j)) = \mathrm{LHS}_k(d) + 1$. If $\Delta_k(d) \ge 1$, then $\Delta_k(d^+(i,j)) \ge 0$. If $d_j < k$, then $\mathrm{RHS}_k(d^+(i,j)) = \mathrm{RHS}_k(d) + 1$ and $\Delta_k(d^+(i,j)) \ge 0$. If $\Delta_k(d) = 0$ and $d_j \ge k$, then $\mathrm{RHS}_k(d^+(i,j)) = \mathrm{RHS}_k(d)$ and hence $\Delta_k\left(d^+(i,j)\right) = -1$. Hence $\Delta_k\left(d^+(i,j)\right) < 0$ is equivalent to condition (2).

We now consider forced non-edges of d. By Theorem 2.1, $\{i,j\}$ is a forced edge if and only if there exists an integer k such that $\Delta_k\left(d^-(i,j)\right) < 0$. We show that this happens if and only if condition (3) or condition (4) holds. Let k be an arbitrary element of $[n]$. Note that if $k \ge i$ then neither (3) nor (4) holds, and $\mathrm{LHS}_k(d^-(i,j)) < \mathrm{LHS}_k(d)$, forcing $\Delta_k(d^-(i,j)) \ge 0$. Assume now that $k < i$. This forces $\mathrm{LHS}_k(d^-(i,j)) = \mathrm{LHS}_k(d)$.

Case: $k \le d_j$. Neither condition (3) nor condition (4) holds, and $\mathrm{RHS}_k(d^-(i,j)) = \mathrm{RHS}_k(d)$, so $\Delta_k\left(d^-(i,j)\right) \ge 0$.

Case: $d_i < k$. Here condition (4) fails. Since $\mathrm{RHS}_k(d^-(i,j)) = \mathrm{RHS}_k(d) - 2$, we see that $\Delta_k\left(d^-(i,j)\right) < 0$ if and only if $\Delta_k(d) \le 1$, which is equivalent to condition (3).

Case: $d_j \le k \le d_i$. Here condition (3) fails. Since $d_j \le k$, we have $\mathrm{RHS}_k(d^-(i,j)) = \mathrm{RHS}_k(d) - 1$. If $\Delta_k(d) \ge 1$, then $\Delta_k(d^-(i,j)) \ge 0$. If $\Delta_k(d) = 0$, then $\Delta_k\left(d^+(i,j)\right) = -1$. Hence $\Delta_k\left(d^+(i,j)\right) < 0$ is equivalent to condition (4). □

3 Structure induced by forced pairs

Theorems 2.1 and 2.3 allow us to determine if a single pair of vertices comprises a forced edge or forced non-edge by examining the degree sequence. In this section we determine the structure of all forcible adjacency relationships by describing the envelope graphs $I(d)$ and $U(d)$ introduced in Section 1.

Recall that the edge set of $I(d)$ is the intersection of all edge sets of realizations of d, and $U(d)$ is the union of all edge sets of realizations, and realizations have the property that vertex i has degree d_i for all $i \in [n]$. Given a degree sequence d and a realization G of d, observe that $I(d) = U(d) = G$ if and only if G is the unique realization of d; by definition this happens if and only if G is a threshold graph. As we will see in Theorem 3.2, threshold graphs have a more general connection to envelope graphs of degree sequences.

Before proceeding we need a few basic definitions and results. An *alternating 4-cycle* in a graph G is a configuration involving four vertices a, b, c, d of G such that ab, cd are edges of G and neither ad nor bc is an edge. Observe that if G has such an alternating 4-cycle, then deleting ab and cd from G and adding edges bc and ad creates another graph in which every vertex has the same degree as it previously had in G. It follows that none of the pairs $\{a,b\}$, $\{b,c\}$, $\{c,d\}$, $\{a,d\}$ is forced in G. By a well known result of Fulkerson, Hoffman, and McAndrew [7], a graph G shares its degree sequence with some other realization if and only if G contains an alternating 4-cycle. Thus threshold graphs are precisely those without alternating 4-cycles.

Lemma 3.1. *Suppose that $d_k \ge d_j$. If ij is a forced edge for d, then ik is also a forced edge. If ik is a forced non-edge, then ij is also a forced non-edge.*

Proof. Suppose that ij is a forced edge. If ik is not a forced edge, then let G be a realization of d where ik is not an edge. Since $d_k \ge d_j$ and j has a neighbor (namely i) that k does not, k must be adjacent to a vertex ℓ to which j is not. However, then there is an alternating 4-cycle with vertices i, j, k, ℓ that contains the edge ij, a contradiction, since ij is a forced edge. By considering complementary graphs and sequences, this argument also shows that if ik is a forced non-edge, then ij is a forced non-edge as well. □

Theorem 3.2. *For any degree sequence d, both $I(d)$ and $U(d)$ are threshold graphs.*

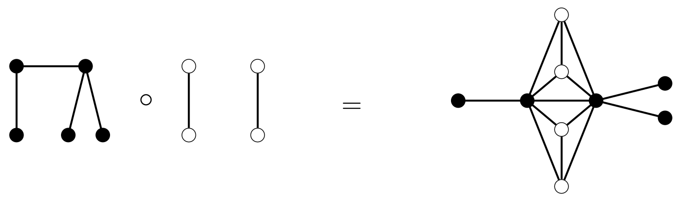


Figure 1: The composition of a split graph and a graph.

Proof. If $I(d)$ is not a threshold graph, then it contains an alternating 4-cycle with edges we denote by pq, rs and non-edges qr, ps. Without loss of generality we may suppose that p has the smallest among the degrees of these four vertices. Since q is forcibly adjacent to p, by Lemma 3.1 it must be forcibly adjacent to r, a contradiction, since qr is not an edge in $I(d)$.

Similarly, if $U(d)$ has an alternating 4-cycle on p, q, r, s as above, and if we assume that p has the largest degree of these vertices, then by Lemma 3.1 since s is forcibly nonadjacent to p it must be forcibly nonadjacent to r, a contradiction. □

We now turn to a precise description of $I(d)$ and $U(d)$. Examining the four scenarios in Theorem 2.3 under which forcible adjacency relationships may occur, we notice that if for some k we have $\Delta_k = 0$, then

- the set $B = \{i : 1 \leq i \leq k\}$ is a clique in which all pairs of vertices are forcibly adjacent;
- the set $A = \{i : i > k \text{ and } d_i < k\}$ is an independent set in which all pairs of vertices are forcibly nonadjacent; and
- each vertex in $C = \{i : i > k \text{ and } d_i \geq k\}$ belongs to a forced edge with each vertex in B and belongs to a forced non-edge with each vertex in A.

This structure of adjacencies within and between A, B, and C has arisen many times in the literature. In particular, R.I. Tyshkevich and others described a graph decomposition based upon it, which we now briefly recall. Our presentation is adapted from [16], which contains a more detailed presentation and references to earlier papers.

A *split graph* is a graph G for which there exist disjoint sets A, B such that A is an independent set and B is a clique in G, and $V(G) = A \cup B$. We define an operation $\circ$ with two inputs. The first input is a split graph F with a given partition of its vertex set into an independent set A and a clique B (denote this by (F, A, B)), and the second is an arbitrary graph H. The *composition* $(F, A, B) \circ H$ is defined as the graph resulting from adding to the disjoint union $F + H$ all edges having an endpoint in each of B and $V(H)$. For example, if we take the composition of the 5-vertex split graph with degree sequence $(3, 2, 1, 1, 1)$ (with the unique partition of its vertex set into a clique and an independent set) and the graph $2K_2$, then the result is the graph on the right in Figure 1.

If G contains nonempty induced subgraphs H and F and a partition A, B of $V(F)$ such that $G = (F, A, B) \circ H$, then G is *(canonically) decomposable*; otherwise G is *indecomposable*. Tyshkevich showed in [16] that each graph can be expressed as a composition $(G_k, A_k, B_k) \circ \cdots \circ (G_1, A_1, B_1) \circ G_0$ of indecomposable induced subgraphs (note that $\circ$ is associative); indecomposable graphs are those for which $k = 0$. This representation is known as the *canonical decomposition* of the graph and is unique up to isomorphism of the indecomposable (partitioned) subgraphs involved.

As observed by Tyshkevich [16], the canonical decomposition corresponds in a natural way with a decomposition of degree sequences of graphs, and it is possible from the degree sequence to deduce whether a graph is canonically indecomposable. In [2], the author made explicit some relationships between the canonical decomosition of degree sequences and the Erdos–Gallai inequalities recalled in Section 2.

Let $EG(d)$ be the list of nonnegative integers ℓ for which $\Delta_\ell = 0$, ordered from smallest to largest. We adopt the convention that empty sums have a value of zero in the definitions of $\mathrm{LHS}_\ell(d)$ and $\mathrm{RHS}_\ell(d)$; thus $\Delta_0(d) = 0$ for all d, and $EG(d)$ always begins with 0.

Theorem 3.3 ([10, 15, 17])**.** *A graph G with degree sequence $d = (d_1, \ldots, d_n)$ is split if and only if $m(d)$ is a term of $EG(d)$.*

Theorem 3.4 ([2], Theorem 5.6)**.** *Let G be a graph with degree sequence $d = (d_1, \ldots, d_n)$ and vertex set $[n]$. Suppose that $(G_k, A_k, B_k) \circ \cdots \circ (G_1, A_1, B_1) \circ G_0$ is the canonical decomposition of G, where A_0 and B_0 partition $V(G_0)$ into an independent set and a clique, respectively, if G_0 is split. A nonempty set $W \subseteq V(G)$ is equal to the clique B_j in the canonical component G_j if and only if $W = \{\ell : t < \ell \leq t'\}$ for a pair t, t' of consecutive terms in $EG(d)$. In this case the corresponding independent set A_j is precisely the set $\{\ell \in [n] : t < d_\ell < t'\}$. Given a term t of $EG(d)$, if $\ell > t$ and $d_\ell = t$ then the canonical component containing ℓ consists of only one vertex.*

Thus the condition $\Delta_k(d) = 0$ in Theorem 2.3 is intimately related to the composition operation $\circ$ and to the canonical decomposition. More generally, we now show that $\Delta_k(d)$ actually measures how far a realization of d is from being a composition of the form described earlier, with slightly relaxed definitions of the sets A, B, and C. Given a subset S of a vertex set of a graph, let $e(S)$ denote the number of edges in the graph having both endpoints in S, and let $\overline{e}(S)$ be the number of pairs of nonadjacent vertices in S. Given another vertex subset T, disjoint from S, let $e(S,T)$ denote the number of edges having an endpoint both in S and in T, and let $\overline{e}(S,T)$ denote the number of pairs of nonadjacent vertices containing a vertex from each of S and T.

Lemma 3.5. *Let G be an arbitrary realization of a degree sequence $d = (d_1, \ldots, d_n)$. Given fixed $k \in [n]$, let $B = \{i : 1 \leq i \leq k\}$, and let A and C be disjoint sets such that $A \cup C = V(G) - B$, each vertex in A has degree at most k, and each vertex in C has degree at least k.*

The kth Erdős–Gallai difference is given by

$$\Delta_k(d) = 2e(A) + 2\overline{e}(B) + e(A,C) + \overline{e}(B,C).$$

Proof. Observe that summing the degrees in B yields $2e(B) + e(A,B) + e(B,C)$, and a similar statement holds for A. Then

$$\begin{aligned}
\Delta_k(d) &= k(k-1) + \sum_{\ell > k} \min\{k, d_\ell\} - \sum_{\ell \leq k} d_\ell \\
&= k(k-1) + \sum_{\ell \in A} d_\ell + k|C| - (2e(B) + e(A,B) + e(B,C)) \\
&= 2\left(\binom{k}{2} - e(B)\right) + 2e(A) + e(A,C) + (|B||C| - e(B,C)),
\end{aligned}$$

and the claim follows. ☐

Observe that Lemma 3.5, besides providing the corollary below, gives another illustration of the role that a value of 0 or 1 for $\Delta_k(d)$ has in producing forced edges and non-edges (as in Theorem 2.3) and in forcing the canonical decomposition structure (as in Theorem 3.4).

Corollary 3.6. *Let $d = (d_1, \ldots, d_n)$ be a degree sequence. For all $k > m(d)$, we have $\Delta_k(d) \geq 2$.*

Proof. Since $k > m(d)$, we know that $d_k < k - 1$, so any set B of k vertices of highest degree in a realization of d cannot form a clique; thus $\Delta_k(d) \geq 2$ by Lemma 3.5. ☐

We now use our results in Section 2 to determine $I(d)$ and $U(d)$. We begin with a quick observation and some definitions we will use throughout the theorem and its proof.

Observation 3.7. *If an indecomposable canonical component (G_i, A_i, B_i) has more than one vertex, then both A_i and B_i must have at least two vertices.*

Let G be a graph with degree sequence $d = (d_1, \ldots, d_n)$ on vertex set $[n]$, and suppose that G has canonical decomposition $(G_k, A_k, B_k) \circ \cdots \circ (G_1, A_1, B_1) \circ G_0$.

Let p be the last element of $EG(d)$, and let q be the largest value of k for which $\Delta_k(d) \leq 1$. If G_0 is split, let A_0, B_0 be a partition of G_0 into an independent set and a clique, respectively. If G_0 is not split, then define

$$\begin{aligned} B_0' &= \{i \in [n] : p < i \leq q\}; \\ A_0' &= V(G_0) - B_0'; \\ A_0'' &= \{i \in [n] : i > q \text{ and } p < d_i < q\}; \\ B_0'' &= V(G_0) - A_0''. \end{aligned}$$

Further let C_1 (respectively C_2) denote an abstract split canonical component consisting of a single vertex lying in the independent set of the component (in the clique of the component). For $i \in \{1, 2\}$ and j a natural number, let C_i^j represent the expression $C_i \circ \cdots \circ C_i$, where there are j terms C_i in the composition.

Theorem 3.8. *Given the graph G with degree sequence d, with the canonical components of G and other sets as defined above, the graph $I(d)$ is isomorphic to*

$$C_1^{|A_k|} \circ C_2^{|B_k|} \circ \cdots \circ C_1^{|A_1|} \circ C_2^{|B_1|} \circ C_1^{|A_0|} \circ C_2^{|B_0|}$$

if G is split, and to

$$C_1^{|A_k|} \circ C_2^{|B_k|} \circ \cdots \circ C_1^{|A_1|} \circ C_2^{|B_1|} \circ C_1^{|A_0'|} \circ C_2^{|B_0'|}$$

otherwise.

Similarly, the graph $U(d)$ is isomorphic to

$$C_2^{|B_k|} \circ C_1^{|A_k|} \circ \cdots \circ C_2^{|B_1|} \circ C_1^{|A_1|} \circ C_2^{|B_0|} \circ C_1^{|A_0|}$$

if G is split, and to

$$C_2^{|B_k|} \circ C_1^{|A_k|} \circ \cdots \circ C_2^{|B_1|} \circ C_1^{|A_1|} \circ C_2^{|B_0''|} \circ C_1^{|A_0''|}$$

otherwise.

Proof. By definition, $\Delta_q \leq 1$, and by Theorem 3.4, it follows that each vertex $i \in [n]$ of G belonging to a set B_j for $j \geq 0$ satisfies $i \leq q$. Theorem 2.3(1) implies that any two vertices in a clique B_j are joined by a forced edge, as are any two vertices in B_0', if G_0 is not split.

Consider any two vertices i, i' belonging to the set A_j for $j \geq 0$. By Observation 3.7, B_j must be nonempty, so it follows from Theorem 3.4 that $d_i < p$ and $d_{i'} < p$. Since i, i' do not belong to B_ℓ for any ℓ, Theorem 3.4 also implies that $i, i' > p$, so by Theorem 2.3(3) the pair $\{i, i'\}$ is a forced non-edge. Similarly, any two vertices in A_0'' are forcibly nonadjacent.

Now suppose that vertices i, i' satisfy $i \in B_j$ and $i' \in V(G_\ell)$, with $\ell < j$. From the adjacencies required by the canonical decomposition we see that $d_{i'}$ is at least as large as $|B_k \cup B_{k-1} \cup \cdots \cup B_j|$, and it follows from Theorem 3.4 that this latter number equals a term t' of $EG(d)$ for which $i \leq t'$. By Theorem 2.3(2), the pair $\{i, i'\}$ is a forced edge.

Suppose instead that vertices i, i' satisfy $i \in A_j$ and $i' \in V(G_\ell)$, with $\ell < j$. Again letting $t' = |B_k \cup B_{k-1} \cup \cdots \cup B_j|$, Theorem 3.4 implies that $\Delta_{t'}(d) = 0$ and that $i, i' > t'$. The adjacencies of the canonical decomposition imply that $d_{i'} \geq t'$ and that $d_i \leq t'$. Theorem 2.3(4) then implies that $\{i, i'\}$ is a forced non-edge.

We now show that all other pairs of vertices in G are unforced, beginning with those within a split canonical component. Suppose that $i \in B_j$ and $i' \in A_j$ for some $j \geq 0$. Any neighbor of i' in G other than i is a neighbor of i; furthermore, since G_j is indecomposable, i' has at least one non-neighbor in B_j, which must be a neighbor of i, so we conclude that $d_i > d_{i'}$ and $i < i'$. Now by Theorem 3.4, there exist consecutive terms t and t' of $EG(d)$ such that $t < i \leq t'$ and $t < d_{i'} < t'$.

We verify that none of the conditions in Theorem 2.3 are satisfied by the pair $\{i, i'\}$. First, since G_j is indecomposable, vertex t' must have at least one neighbor in A_j, so $d_{t'} \geq t'$. Thus $i' > t'$, and since $d_{i'} < t'$, we see that $\{1, \dots, i'\}$ is not a clique, so by Lemma 3.5 we see that $\Delta_\ell(d) \geq 2$ for all $\ell \geq i'$. Thus condition (1) of Theorem 2.3 does not apply to the pair $\{i, i'\}$.

Condition (2) does not apply, since t' is the smallest term of $EG(d)$ at least as large as i, and $d_{i'} < t'$. Condition (3) likewise cannot apply, since $\{1, \dots, t'\}$ is a clique and hence $d_i \geq i - 1$. Finally, since t' is the smallest term of $EG(d)$ at least as large as $d_{i'}$, but $i \leq t'$, condition (4) does not apply.

It remains to show that $\{i, i'\}$ is unforced if G_0 is not split and vertices $i, i' \in V(G_0)$ don't both belong to B'_0 or both belong to A''_0. Assume that $i < i'$.

If at least one of i, i' does not belong to B'_0, then we claim that $\{i, i'\}$ cannot be a forced edge. Indeed, note that $i' > q$ and $i > p$, so neither of conditions (1) or (2) of Theorem 2.3 applies.

If at least one of i, i' does not belong to A''_0, then we claim that $\{i, i'\}$ is not a forced non-edge. Indeed, note that $d_i \geq d_{i'}$, and since G_0 is indecomposable and has more than one vertex, we have $d_{i'} > p$; this implies that $\{i, i'\}$ fails condition (4). We also see that $i \leq q$ or $d_i \geq q$; in either case condition (3) does not apply.

Having characterized all pairs of vertices as forced or unforced, we can now summarize the structure of $I(d)$ and $U(d)$. If we form a correspondence between each vertex in A_j (respectively, in A'_0, in A''_0, in B_j, in B'_0, in B''_0) with a vertex of $C_1^{|A_j|}$ (of $C_1^{|A'_0|}$, of $C_1^{|A''_0|}$, of $C_2^{|B_j|}$, of $C_2^{|B'_0|}$, of $C_2^{|B''_0|}$) in the claimed expressions for $I(d)$ and $U(d)$, the correspondence naturally leads to an exact correspondence between the edges in either of the first two expressions and the edges in $I(d)$. Likewise, the edges in the third and fourth expressions in the theorem statement correspond precisely to the edges in $U(d)$. $\square$

A well known and useful characterization of threshold graphs (see [11, Theorem 1.2.4]) states that a graph is threshold if and only if it can be constructed from a single vertex by iteratively adding dominating and/or isolated vertices. The expressions in Theorem 3.8 describe how the envelope graphs of d can be constructed in this way: because of the requirements of the operation $\circ$, as we read from right to left, a term C_1^a corresponds to adding a isolated vertices in sequence, and a term C_2^b corresponds to adding b dominating vertices.

Example. If d is the degree sequence of the graph on the right in Figure 1, then $I(d)$ is formed by starting with a single vertex, adding three more isolated vertices in turn, adding two dominating vertices, and finishing with three more isolated vertices. The graph $U(d)$ is formed by starting with a single vertex, adding three more dominating vertices, three more isolated vertices, and then two more dominating vertices.

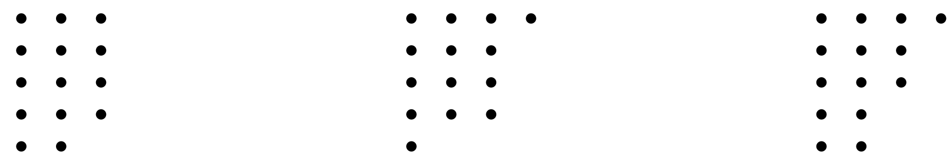
Note that if d is threshold, then $I(d) = U(d) = d$ (as expected), because every canonical component (G_i, A_i, B_i) of a threshold graph contains only a single vertex (we see this from the dominating/isolated vertex construction described above), so for all i either A_i or B_i is empty, and the expressions in Theorem 3.8 simplify to return the canonical decomposition of the unique realization of d.

4 Threshold graphs and the dominance order

In this section we compare the forcible adjacency relationships of degree sequences that are comparable under the dominance order.

Given lists $a = (a_1, \dots, a_k)$ and $b = (b_1, \dots, b_\ell)$ of nonnegative integers, with $a_1 \geq \cdots \geq a_k$ and $b_1 \geq \cdots \geq b_\ell$, we say that *a majorizes* b and write $a \succeq b$ if $\sum_{i=1}^k a_i = \sum_{i=1}^\ell b_i$ and for each $j \in \{1, \dots, \min(k, \ell)\}$ we have $\sum_{i=1}^j a_i \geq \sum_{i=1}^j b_i$. The partial order induced by $\succeq$ on lists of nonnegative integers with a fixed sum s and length n is called the *dominance* (or *majorization*) *order*, and we denote the associated partially ordered set by $\mathcal{P}_{s,n}$.

(We remark that requiring sequences to have the same length and allowing terms to equal 0 are slight departures from how the dominance poset is often described. We do so here for convenience in the statements of results below.)



$(3,3,3,3,2)$ $(4,3,3,3,1)$ $(4,3,3,2,2)$

Figure 2: Ferrers diagrams depicting elementary transformations

The dominance order plays an important role in the study of graphic lists. It is known that if for $a, b \in \mathcal{P}_{s,n}$ it is true that a is graphic and $a \succeq b$, then b is also graphic; thus the degree sequences form an ideal in $\mathcal{P}_{s,n}$. The maximal graphic lists are precisely the threshold sequences [14].

We define a *unit transformation* on a nonincreasing integer sequence to be the act of decreasing a sequence term by 1 and increasing an earlier term by 1 while maintaining the descending order of terms. This operation is best illustrated by the Ferrers diagram of the sequences, where sequence terms are depicted by left-justified rows of dots. Note that if a results from a unit tranformation on b, then the Ferrers diagram of a differs from that of b by the removal of a dot from one row of b to a row higher up in the diagram. In Figure 2, the second and third sequences each result from a unit transformation on the first sequence.

A fundamental lemma due to Muirhead [13] says that $a \succeq b$ if and only if a may be obtained by performing a sequence of unit transformations on b.

Theorem 4.1. *Let d and e be graphic elements of $\mathcal{P}_{s,n}$. If $d \succeq e$ and $\{i,j\}$ is a forced pair for e, then $\{i,j\}$ is a forced pair for d.*

Proof. We may obtain d from a sequence of unit transformations on the sequence e. The intermediate sequences resulting from these individual transformations all majorize e, so it suffices to assume that d can be obtained from just one unit transformation. In other words, we assume that there exist indices s and t such that $s < t$ and

$$d_\ell = \begin{cases} e_\ell + 1 & \ell = s \\ e_\ell - 1 & \ell = t \\ e_\ell & \text{otherwise} \end{cases}.$$

Suppose now that $\{i,j\}$ is a forced edge for e. By Theorem 2.1 $e^+(i,j)$ is not graphic, so by Theorem 2.2 there exists an index k such that $k \geq i$ and

$$\sum_{\ell \leq k} e^+(i,j)_\ell > k(k-1) + \sum_{\ell > k} \min\{k, e^+(i,j)_\ell\}.$$

Since the actions of increasing two terms of a sequence and performing a unit transformation on a sequence together yield the same result regardless of the order in which they are carried out, we have

$$\begin{aligned} \sum_{\ell \leq k} d^+(i,j)_\ell &\geq \sum_{\ell \leq k} e^+(i,j)_\ell \\ &> k(k-1) + \sum_{\ell > k} \min\{k, e^+(i,j)_\ell\} \\ &\geq k(k-1) + \sum_{\ell > k} \min\{k, d^+(i,j)_\ell\}. \end{aligned}$$

Thus $d^+(i,j)$ is not graphic, and by Theorem 2.1 $\{i,j\}$ is a forced edge for d.

A similar argument holds if $\{i,j\}$ is a forced non-edge for e, making $\{i,j\}$ a forced non-edge for d. $\square$

Example. The degree sequence $(2,1,1,1,1)$ is majorized by $(2,2,1,1,0)$, which is in turn majorized by $(3,1,1,1,0)$. The first sequence has has no forced pair. In the second sequence vertex 5 is forcibly nonadjacent to all other vertices, $\{3,4\}$ is a forced non-edge, and $\{1,2\}$ is a forced edge. These relationships are all preserved in $(3,1,1,1,0)$, and every other pair of vertices is forced as well, since $(3,1,1,1,0)$ is a threshold sequence.

As illustrated in the previous example, forcible adjacency relationships come into existence as we progress upward in the dominance order, and they persist until the threshold sequences are reached, where every pair of vertices is a forced pair. Thus the proportion of all vertex pairs that are forced may be considered a measure of how close a degree sequence is to being a threshold sequence.

Our results now yield a consequence of Merris [12, Lemma 3.3]. We call a degree sequence *split* if it has a realization that is a split graph.

Corollary 4.2. *Let d and e be graphic elements of $\mathcal{P}_{s,n}$. If $d \succeq e$ and e is split, then d is split.*

Proof. As usual, let G be a realization of e with vertex set $[n]$. Since e is split, by Theorem 3.3 we know that $\Delta_{m(e)} = 0$. By Theorem 2.3 every pair of vertices from $\{1,\ldots,m(e)\}$ forms a forced edge. Likewise, any pair of vertices from $\{m(e)+1,\ldots,n\}$ forms a forced non-edge. By Theorem 4.1, these forcible adjacency relationships exist also for d, so $\{1,\ldots,m(e)\},\{m(e)+1,\ldots,n\}$ is a partition of the vertex set of any realization of d into a clique and an independent set; hence d is also split. $\square$

Note that by Theorems 2.3 and 3.4, adjacencies and non-adjacencies between vertices in distinct canonical components, as well as adjacencies between two clique vertices and non-adjacencies between two independent-set vertices in split canonical components, are all forcible adjacency relationships. Thus every realization of the degree sequence of a canonically decomposable graph is canonically decomposable. It is natural to then, as we did for split sequences, refer to a degree sequence itself as *decomposable* if it has a decomposable realization.

The forcible adjacency relationships between canonical components and inside split components further imply, via an argument similar to that of Corollary 4.2, that canonically decoposable graphs have the same majorization property that split graphs do.

Corollary 4.3. *Let d and e be graphic elements of $\mathcal{P}_{s,n}$. If $d \succeq e$ and e is canonically decomposable, then d is canonically decomposable.*

More generally, all sequences with at least one forced pair form an upward-closed set in $\mathcal{P}_{s,n}$. We now show, in fact that these sequences lie close in $\mathcal{P}_{s,n}$ to split or decomposable sequences. The key will be the observation that according to Lemma 3.5, having a small Erdős–Gallai difference requires a graph to have a vertex partition that closely resembles that of a split or decomposable graph.

Our measurement of "closeness" in $\mathcal{P}_{s,n}$ will involve counting covering relationships. A unit transformation on a nonincreasing integer sequence is said to be an *elementary transformation* if there is no longer sequence of unit transformations that produces the same result; in other words, an elementary transformation changes an integer sequence into one that immediately covers it in $\mathcal{P}_{s,n}$. As shown by Brylawski [3], a unit transformation on a sequence $b = (b_1,\ldots,b_\ell)$ is an elementary transformation if and only if, supposing that the pth term of b is increased and the qth term is decreased, we have either $q = p+1$ or $b_p = b_q$. The rightmost sequence in Figure 2 shows the result of an elementary transformation on the original sequence, while the middle sequence shows a non-elementary unit transformation.

Theorem 4.4. *If e is a graphic sequence in $\mathcal{P}_{s,n}$ that induces any forcible adjacency relationships among the vertices of its realizations, then some sequence d that is split or canonically decomposable is located at most three elementary transformations above e in $\mathcal{P}_{s,n}$.*

Proof. By Theorem 2.3, e can only force vertices to be adjacent or nonadjacent if $\Delta_k(e) \le 1$ for some positive k. If for such a k we have $\Delta_k(e) = 0$, then by Theorems 3.3 and 3.4 we may let $d = e$.

Suppose instead that $\Delta_k(e) = 1$ for some k, and let G be a realization of e on vertex set $[n]$. Partition $V(G)$ into sets A, B, and C as in the statement of Lemma 3.5, with $B = \{i : 1 \le i \le k\}$. Since $\Delta_k(e) = 1$, this lemma implies that A is an independent set, B is a clique, and exactly one of the following cases holds:

(1) there is a single edge joining a vertex in A with a vertex in C, and all edges possible exist joining vertices in B with vertices in C;

(2) there is a single non-edge between a vertex in B and a vertex in C, and there are no edges joining vertices in A with vertices in C.

We consider each of these cases in turn. We will use the following statement, which is proved using elementary edge-switching arguments:

FACT [1, Lemma 3.2]: *Given a vertex v of a graph G and a set $T \subseteq V(G) - \{v\}$, suppose that v has p neighbors in T. For any set S of p vertices of T having the highest degrees in G, there exists a graph G' with the same vertex set as G in which the neighborhood of v, restricted to T, is S, all neighbors of v outside of T are the same as they are in G, and every vertex has the same degree in G' as in G.*

In the first case, let a be the vertex of A having a neighbor in C. By the fact above we may assume that the neighbor of a in C is a vertex having the highest degree in G among vertices of C; call this neighbor c. Since the degree of a is at most k, there must be some vertex in B to which a is not adjacent; using the fact again, we may assume that this non-neighbor (call it b) has the smallest degree in G among vertices of B. Now deleting edge ac and adding edge ab produces a graph having a degree sequence d which, using partition A, B, C, we see is canonically decomposable.

In the second case, some vertex of C has a non-neighbor in B. By the fact above, we may assume that the non-neighbor in B is a vertex b having the lowest degree in G among vertices of B. Now since every vertex of C has degree at least k and no neighbors in A, but some vertex in C has a non-neighbor in B, this vertex in C must have a neighbor in C. Using the fact again, we may then assume that the non-neighbor of b in C is a vertex u having smallest degree among the vertices in C. Using the fact once again, we may assume that the neighbors of u in $C - \{u\}$ have as high of degree as possible. Now let c be a neighbor of u that has maximum degree among the vertices of C. Deleting the edge uc and adding the edge ub produces a graph a graph having degree sequence d for which, using partition A, B, C, we see is canonically decomposable.

In both cases, the effect of creating degree sequence d from e was to perform a unit transformation which reduced the largest degree of a vertex in C and increased the smallest degree of a vertex in B. Since by assumption the degrees of vertices in B are the highest in the graph, and the degrees of vertices in C may be assumed to precede those of vertices in A in the degree sequence, the creation of d from e is equivalent to a unit transformation on e. In fact, it is equivalent to an elementary transformation if b and c are the unique vertices with their respective degrees or if b has the same degree as c. Otherwise, we may accomplish this unit transformation using two or three elementary transformations, as follows (we use $\deg(v)$ to denote the degree of a vertex v):

If both $\deg(b)$ and $\deg(c)$ appear multiple times in d and $\deg(b) > \deg(c)+1$, then decrease the last term equal to $\deg(b)$ while increasing the first term equal to $\deg(b)$; decrease the last term equal to $\deg(c)$ while increasing the first term equal to $\deg(c)$; and decrease the term currently equal to $\deg(c)+1$ while increasing the term currently equal to $\deg(b) - 1$.

If $\deg(b)$ appears multiple times in d and either $\deg(c)$ appears only once or $\deg(b) = \deg(c) + 1$, then decrease the last term equal to $\deg(b)$ while increasing the first term equal to $\deg(b)$; then decrease the last term equal to $\deg(c)$ while increasing the first term currently equal to $\deg(b) - 1$.

If $\deg(c)$ appears multiple times in d and $\deg(b)$ appears only once, then decrease the last term equal to $\deg(c)$ while increasing the first term equal to $\deg(c)$; then decrease the last term currently equal to $\deg(c)+1$ while increasing the first term currently equal to $\deg(b)$. □

We conclude by showing that the bound in Theorem 4.4 is sharp for infinitely many degree sequences.

Example. Consider the sequence $s = \big((15+2j)^{(5)}, 6^{(7+2j)}, 3^{(7)}\big)$, where j is any nonnegative integer; note that $\Delta_5(s) = 1$ and that $\Delta_k(s) \neq 0$ for all positive k. Let s' and s'' denote sequences obtained by performing respectively one and two elementary transformations on s.

Observe by inspection that $s''_7 \geq s'_7 \geq s_7 = 6$ and $s''_8 = s'_8 = s_8 < 7$; thus $m(s) = m(s') = m(s'') = 7$. To test whether any of s, s', s'' has a realization that is decomposable, by Theorem 3.4 and Corollary 3.6, it

suffices to test whether equality holds in any of the first seven Erdős–Gallai inequalities for the corresponding sequence.

Recalling our notation from before, we see that since $\mathrm{LHS}_k(s) \leq \mathrm{LHS}_k(s') \leq \mathrm{LHS}_k(s'')$ and $\mathrm{RHS}_k(s) \leq \mathrm{RHS}_k(s') \leq \mathrm{RHS}_k(s'')$, if any of s, s', s'' satisfied the kth Erdős–Gallai inequality with equality, it would follow that $\mathrm{RHS}_k(s) \leq LHS_k(s'')$. Now consider the table below, which shows the maximum possible value for $\mathrm{LHS}_k(s'')$ and the value of $\mathrm{RHS}_k(s)$ for each $k \in \{1, \ldots, 7\}$.

k	max $\mathrm{LHS}_k(s'')$	$\mathrm{RHS}_k(s)$
1	$16 + 2j$	$18 + 2j$
2	$32 + 4j$	$36 + 4j$
3	$47 + 6j$	$54 + 6j$
4	$61 + 8j$	$65 + 8j$
5	$75 + 10j$	$76 + 10j$
6	$82 + 10j$	$87 + 12j$
7	$89 + 10j$	$93 + 12j$

We see that each of s, s', s'' is graphic. Furthermore, since in no case does $\mathrm{RHS}_k(s) \leq LHS_k(s'')$, we conclude that any canonically decomposable degree sequence that majorizes s must be separated from s by at least three elementary transformations. (As guaranteed above, the sequence $\left(16 + 2j, (15 + 2j)^{(4)}, 6^{(6+2j)}, 5, 3^{(7)}\right)$ is located three elementary transformations above s and is the degree sequence of a canonically decomposable graph.)

References

[1] M.D. Barrus and E. Donovan, Neighborhood degree lists of graphs, submitted.

[2] M.D. Barrus, Hereditary unigraphs and Erdos–Gallai equalities, Discrete Math., 313 (2013), no. 21, 2469–2481.

[3] T. Brylawski, The lattice of integer partitions, Discrete Math., 6 (1973), no. 3, 201–219.

[4] V. Chvátal and P.L. Hammer, Set-packing and threshold graphs, Research Report, Comp. Sci. Dept. University of Waterloo, Canada CORR 73-21 (1973).

[5] V. Chvátal and P.L. Hammer, Aggregation of inequalities in integer programming. In P.L. Hammer, E.L. Johnson, B.H. Korte, and G.L. Nemhauser, editors, Studies in Integer Programming, pages 145–162. North-Holland, New York, 1977. Annals of Discrete Mathematics, 1.

[6] P. Erdős and T. Gallai, Graphen mit Punkten vorgeschriebenen Grades, Math. Lapok., 11 (1960), 264–272.

[7] D.R. Fulkerson, A.J. Hoffman, and M.H. McAndrew, Some properties of graphs with multiple edges, Can. J. Math., 17 (1965), 166177.

[8] P.L. Hammer, T. Ibaraki, and B. Simeone, Degree sequences of threshold graphs, Congres. Numer. 21 (1978) 329–355.

[9] P.L. Hammer, T. Ibaraki, and B. Simeone, Threshold sequences, SIAM J. Algebraic Discrete Methods 2 (1981) 39–49.

[10] P.L. Hammer and B. Simeone, The splittance of a graph, Combinatorica 1 (1981) 275–284.

[11] N.V.R. Mahadev and U.N. Peled, Threshold graphs and related topics. Annals of Discrete Mathematics, 56. North-Holland Publishing Co., Amsterdam, 1995.

[12] R. Merris, Split graphs, European J. Combin. 24 (2003), no. 4, 413–430.

[13] R.F. Muirhead, Some methods applicable to identities and inequalities of symmetric algebraic functions on n letters, Proceedings of the Edinburgh Mathematical Society, 21 (1903), 144–157.

[14] U.N. Peled and M.K. Srinivasan, The polytope of degree sequences, Linear Algebra Appl., 114/115 (1989), 349–377.

[15] R.I. Tyshkevich, The canonical decomposition of a graph, Doklady Akad. Nauk, BSSR 24 (1980) 677–679. In Russian.

[16] R. Tyshkevich, Decomposition of graphical sequences and unigraphs, Discrete Math. 220 (2000), no. 1–3, 201–238.

[17] R.I. Tyshkevich, O.I. Melnikow, and V.M. Kotov, On graphs and degree sequences: the canonical decomposition, Kibernetica 6 (1981) 5–8. In Russian.